

Figure 1

Sequence Name:

A.ctg12831-000000.10.0

Figure 1A:

MGKDFMSKTPKAMATKAKIDKWDLIKLSFCTAKETTIRVNRQLTEWEKIFATYSFDKGL
ISRIYNELKQIYKKKTKNPIKKWVKDMNRHFSKEGIYAAKHKMKYSSSLAIREMQIKTT
MRYHLTPVRMAI IKKSGNNRDMDEAGNHHSQQTITRTKNQTPHVLTHRWILQQSHWVTVL
SDISELMHKTDRIVNLMLCMYLLTVDLDRNLDDAKRYSCTPRNYSVNIREELKLANVVF
PRCLLVQRCGNGCGTVNWRSTCNSGKTVKKYHEVLQFEPGHIKRRGRAKTMALVDIQ
LDHHERCDCICSSRPPR

Figure 1B:

GTATATGTAAGAAAGCCTCATCTTTTGATTTTTAATATACAAGATGCTTTCTTTAAGAGA
GCAAGATTCAAAATTTGTTTGTGTTTCAAAATTTAAAAATAAAATTTATCTCCTAAATTTT
CTAAAGACATGTTTTCATATATTTGACCATCCCCTATTTTTGGCAAAGGATTTTAAGAGTCT
AACTCAAAACATATGTAAGCTCTGGTGACCTGGTTATATATACCAAAAAAACATTTGAT
CTATATACACATAGACATGAATATATTTCTGTGTGTGTTTGTGCATATATAACCTCAAAC
ACTATTATTAAATGCAATCCTATATTCTTAGGTATAGAAGTTGATGATATACCTTTCTAC
TTGCCATGGCATTAAACAAAGCAAGGCTGAGACTCAGCAACCACTGTGTTCATTGCATTG
CAGGCTAGTAGTAAGTTTGGTTGCTGGTAGGAAAAGGGTCTCTTATCTCACCTCCTTAA
ACTAAAGGTTCTTTAGGCTTAATGTAAGGATGTGCACATTCTCTTATCGAGGTGGTCTT
GAGCTGCAGATACAATCACATCGTTTCATGGTGATCCAACCTGGATGTCAACTAGAGCCATG
GTCTTAGCTCTACCCCTCCTCTTGATGTGGCCAGGCTCAAACCTGTAATACCTAGGACAAG
AAGCACATCTCCTGTTAGAAAAGCCTTTGGAGTTCAACTCAGTCAGATGCCACCTACTTAT
TACCTTTTTTGACAACTAGTTCTTAGCCCTTTGAGAACCCAAACAGAAGCTATGGGCTTGCT
ATTAGAATGCACACGTTGCTATTAGAATGTACACATTTTTCAAATAATTGACTCCCTGAA
GTGGAGGAATCAATTGATCCAGAGTAATGCCAGCATAACTTACCTGAAGTACCCAGAT
GATTTTCATGTGTCTTAGCAGGTATTTATTAATAGCTTTCTAAGGGCCTGCTTTGGGCCAA
GTACTGTTCCAAATATTATCGTAAAGATCCTTCTGACCAAGGCATGTGTTATAGATGAAT
ACAATACTTGAGCATATTATAGCATGGAGAGGAAAAATGAATACAACCATGAATAAAATA
TGCTGGTATATCTAAATCTTTGGTTGAAGTAAACATGTTGCCCTGGAGTTGCTGGCAAG
ATGGCCGAACAGGAACAGCTCTGGTCTGCAAGTTCCAGCGAGATCAATGCAGAAGGCGGG
TGATTTCTCCATTCCCAACTGAGGTACCCAGTTTCATCTCACTGGGACTGGTTAGACATTG
GGTGCGAGCCACGGAAGGTGAGCTGAAGCAGGGTGGGGTGTCCCCTCAGCCGCGAAGTGC
AAGGGGGTGGGGGATCTCCTTCCCCCAGCCAAGGGAAGCCATGAGAGACTGTACCAGGAG
GAATGGTGCACTCTAGTCCAGATACTGCACTTTTCCCATAGTCTTTGCAACTGGCAGACC
AGGAGATTTCCCCCAGTGCTATGCCACCAGGGCCCTGGGTTTCAAGCACAAAACCTGGGC
GGCCATTGACAGACACCGAGCTAGCCGAGCAGTTTATTTTTCATACCCAGTGGCGC
CTGGAATGCCAGCAAGACAGAACCATTCCTCCAGGGATCCAAGTGGTCTGGCTCAGTGG
GTCCACCCCCATGGAGCCAGCTAGCTAAGATCCACTGGCTTGAAATTCTCCTGCCAGC
ACAGCAGTCTGAGATTGACCTGGGATGCTTGAGCTTGGTGAGGGGAGGGGCGTCTGCCAT
TGCTGAGGCTTGAGTAGGCGAGGCGGTTTACCCTCAAAGTGTAACAAAAGCTACTGGGA
AGTTTGAATGGGGCGCCACCGCAGCTCAGCAAGGCCGCTGTGGCAAACCTGCCTCTCTAG
ATTCTCTCTTTTTGGGCAGGTCTCTCTGAAAGAAAGGCAGCAGCCCCAGTCAGGGACTT
ATAGATAAAACCCCATCTCCTTGGGACAGAGAACCTGGGGGAAGGGGTGGCTGTGGGTG
CAGCTTCTCAGACTTAAACATTCTGCTGGAGGCTCTGAAGAGAGCAGCGGATCTCCCA
GCACAGCATTTGAGCTCTGATAAGGGACAGGCTGCCTCCTCAAGTGGGTCCCTGACCCCC
ATGTATCTGACTGGGAGACATCTCCATTAGGGGCCAATAGACATTTCATACAGGAGAC
AGGGTCTGGAGTGACCTCCAGCAAACCTCAGCAGACCTGCAGCAGAGCGGCTGACTGT
TAGAAGGAAAGTAACAAACAGAAAGGAATAGTATCAACATTAACAAAAAGGACATCCAC
TCAGAGACCCCATCTGAAGGTCAACAACATCAAAGACCAAAGGTAAATAAAACCAAAAAG

T

TAGATGCTATAAATGTGCCAGAATCAGAGATAGGTATAGGCTTGTGCCATTCAACAAGGT
AGTCACGGCCTTGTGTGGTAATTTAAATTTCAATTAGTTAAAACTAAATAAAATTTAAAA
TTTAGTTTTCTATTTGTGCTACACACATTTCAAGTGCCCAACAGCCACATGTTGCTAGTGA
CTACCATATGGAACATTGCAAAATATAGGTTATTCCTATCACTACAGGAAGATCTATTAGA
CAGTGCAGGTGGTTAGGTCAAGCATCCTGAGAAAATTATGAGAGGAAGAACAGAGGAAAT
TAACATCTATCTTTATCTCTTTGATGCTTCCTATGTTTTTGGCTGTTGACTAAATGGATA
CAGCCAGGGCACCAATCATGAAAAACAGTTGAGTAATAATTTAATAGGTCACTATACTT
TTAGGAAATATCCTCTTCCTTATTATACACTATCAAGAGAAGAGAAGTGAATAATAGT
TTCTTCTAATTGTCTACACAATACGTTTTTCTGGAATCCTCCCTTTAACAAAAATCACTA
CCCAAATTATCTTTTATATTCTATGATGGGGTTCAGGACACTGTACTCCCAAATGTTTTA
AGCTGAAGGAATTTGAGAAAAACAAGAAAAAGCAGAAAAGATCACTCTGACCTTCCCTCACC
CTCCATCCTGAAGGAAGTCATAAAACCTAGGATTTTCTGACCTTCCCATGTAGCAAGTCA
TAAGACCTCATGCGAGAGGTGCCCTGTTATACCCAGAAGAAAAGAAGATTCCCCCCCCG
CCTTTTTGTTGAGATGTGGTCTTACTCTGTCAACCAGGCTAGAGTGCAGTTGCATGATC
ATGTTCTGCTGCAACCTTGAACCTCTGGGCTCAAGCAATTCTCCTGCCTCAGCCTCCAA
GTAGCTCAGACCACAGACACATCTCCACACTTGGCTACTTAAAAATAATTCTTTTTTTT
TATTATTTATTTATTTTTTTTTTAGAGATGGGGGGTCTCGCTATGTTGACCCGTTGAT
CTTGAACCTCATAGCCTCAAGGGATCATCCACCTTGGCCTCTCAAAGTGCTGAGATTACA
GGCATGAGCCACTGCTCCTGGCCTAAGAATCCTTATCTCCAAAGACAAAGGTAGAACAAA
TAAGAATCTGAACAAACAGGCCTTGCTAATTTCCCCAGTTTATTACCATTAGATCATAC
TCTGCCCTATCATATTTCTCCCAACTATCCACACTTTATCAAACCTTACTGTAAAAAATT
ATCAGGTTGAACCACTTTTTTGGGTCTTCCTTACCAAGGCCTCTGTGTACGTAAAAAT
ATTCAATAAATGAGTACACTTTTCTTTGTTATTCTGTCTTTTGTATAGTGGCCTCAGC
CATGAACCTAGGAAGGGTGAAGAAAAGGTATTTTCTACTCTATATCTAGTATATTCC
AAGAAGCATAGAATGATTGTTTAGAAAAGAGGAAAAATACTTCAACAGGGTACCTACAAAC
TTTTGAGAGCAAGAAATAATAAAGACAGTCACATTTTACTGCACAGTGAAAACCGCTCAA
AACCTTACAGGACAATAATGACCTGTGAAAAAAGCAATATTAGCTTCGTTAGAAGCT
TTCCAGTTTCCCAGATTTAAATTCCTTAAGGTTGAGCACTGAAATATTTGTGGAGAACTA
CAACAGCTGTATGTGACAAGGTCTCATTTACCCCTTAAGGGTGCTCATCTTTCACAGA
AACTCATTTTACATCTCATCAGTTGCAACCATGAGCTGGGGAAAAGTGGTTTCTCATTTCT
CAGAGGTAAGAATACAGACACAGAATTCAGAATAAAGGACTAGAATTTTTCCCTAATGAG
ACTAGGCTCAAGCAACCCCCACTTCTGGAAATGGAATAAGCCTTTTGTGTTTCCACAG
GGCACTTGACAGTGATAATTGTGACACATCTCAAAGGTATTTAGGAGTCTCAGGGGCAT
ATGACAGGAGTGTGTTTCCAGGTAAGCAGTTACTGGCTGGTGATGGATGTTTATTTTTT
CTTTCATCCAGCACTTGACAGAGTGATTTTATAGAGCTTTCAGAGTCTAGAAAGTAAAA
ATATTAGGACACATGTCTTAAAGAGATGCTTTAGATACAGTTTCAAAGGATGCCAATGTT
TCAAATCCACTCATGCGACTGAACAAAGAGATCAAGAGTTTCCCTCTACACATTAACT
GAGAACAGAAATTTGCAAGTAATGGTTATTGTGTAGAGAGATAAGAAAAGAACTAAGAGG
AGGAGGGTAGCCAGGCCTTCAGTGTTCTTCAAAACTAACTGGCTGGGAGGCAAGGGTCTT
TGGGGGTCTATGGTTCACTCTGCTCTTCCCAATAAGTAATTTTTTAAATCAAGCTAAA
AATTCTATCTGAGCATAGAGTCAAGATAAAAAATAAAGCACTCCACAAATACATCAAAGTC
CTGACCTTGAATGTTCTTCTCACAGTTACTACTAGGTAGAGACCTTGTCACTTCTTAGA
TTATTTTAACTATAACCTGATTACTGCATTCTGCATTTATCCTTGGGGTATGATTAAAGG
TTATAAAATTCAGGTTCTTAGAAGTGAAGGACAACATCCAATCTAGAGTTCATCAGTTCA
TACAAAAATAAAACACAATGAACATAGATGATAGCATTTCTTAGATATGAGAACATTTTT
TTTCTTTGAAAACAGTGGGGTGGATCAAGCAATCTTATGACTTAAGAAAATATGTTCC
ATGTAGGGCTAAAAATAGTGTTCCTGTCTGAAGTAAGTATGGTCATGGTGAACAATTT
CCTACATTTTCCAAGTTGAAAATATCAATTATTACATACTCATTTTTTGCTTAATCCAGAA
GTTACCTCCTAACTTACTACGATTTTTTGTGTTGTTTTTCAAGATAGAAATCAAAACCAAAA
CACCACCACCACCACAAAAGCCTCCAAAAAACAGTTTTTACAAAGCTGGAGATTTTGTTTA
GAGAGAACTACCTATTCCTTTGTTGTGGTGGACACTTTTTTAAAAATTATTATTACTTT
AAGTTGTGGGATACTTGTGCAGAACGTGCAGGTTTGTACGTAGGTATACACGTGCCATG
GTGGTTTGTGTCACCTTTAACCCATCACCTACATTAGGTATTTCTCCTAATGCTATCCC
TCCCCAACCCCCAACCCACCGACAGCCCCCAGTGTGTGATGTTCCCTTCCCTGTGTCCC
TGTGTTCAACTCTCACTTATGAGTGAGAACATGTGGTGGTTGGTTTTCTGTTCCTGTGTT
AGTTTGCTGAGAATGATGGTTTCTAGCTTCATCCGTGTCTCTGCAAAGGACGTGAACCTCA
TCCTTTTTTAAAGGCTGCAAAGTATTCCATGGTGTATATGTGCCACATTTTCTTTATCCAG

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

TCTATCACTGATGGGCATTTGGGTTGGTTCCAAGTCTTTGCTATTGTGAATAGTGCTGCA
ATAAACATATGTGTGCATGTGTCTTTATAGCAGAATGATCTATAATCATTTGGTATATAC
CCAGTAATGGGATTGCTGTGGACACTGATGGAGAGCAACTCAAACAGTAATTCCTAACAA
TGATGTTTGTCTTCATATTTTAGTTTATATTTTAAAAACATATTTCTGCATGATGTTTAA
TCTTGCAAAATATATTTTCATATGCATTATTTTATTTGACCCTCAGAGCAACTCTGGAAGG
GGTTATCTGATATATTAATTTTTCAGTCCTATTTAAATATGTGAAAACCTGAGTCAAACT
TGTAGTACTTGCATGTGTTTCCACAATAATAAAGTAATTAGAGCTGCTGCCTATGTCACG
GAAGAGCTCTAACTCCTCTTCAGTCTACTCTCTCTCTCTCTCTATGTGTGTGTGTCTG
TGTATATATATATATATATATACACAGACACACACACACATATATATATATATATATA
TGTAATTTTTTTTTGAGATGGAGTTTCTTTTACCCAGGCTGGAGTGCAATGGCGCGATC
TCAGCTCACTGCAACCTCCGCTCCCGGGTTCAAGTGATTCTCCCGCTCAGCCTCCTGA
GTAGCTGGGATTACAGGTGCCTGCCATCACGCCTGGCTAATTTTTGTATTTTAGTAGAG
ACGGGGTTTTGCCATGTTGGCCACGCTGTTCTTGAACCTCTGACCTCAGGTGATCCACCC
ACCTCGGCCTCCCAAAGTGCTAGGATTACAGGCGTGAGCCACTGCGCCCGCCTATCTTT
TATGCTCAATAAAGTGCTATATATTACTAGTGATAAGAAAAACCATAGACCTGGAACCA
GTGGATTGGAGTTGGAGTCTGGCTCTGTGATTTACTGACCAGGCCATGCTGGCCAACCT
ACTGATCCTCTTAGGGCTCAGTTTACTCACCTATAAAAAAGAATAAAAAAGTGACATT
TTTACAATATTGTTGATGGGGTGCTAAATGAGCTAATGATTACTTACAATAGTTTTGTCT
AAGCAATAAATGCCATTGTTACTATTATTAAGGTTATTGTGCTCTTTAAGTTTTTGTTC
TCCTATATTCTTTTACTTTTTGTCTATTTTTGCATAATCCAAGAGCTGATTCCCTGAGTAG
CACTTAATAATAGTTAAAAATAAACTGAGCATTTCCGGTTGAAACATACTTCATATTTAGA
AAATAAAAAAGATCTTTGACAATTTATATTTTGCTATCTTTATTGTGCCAAATTTATTTAA
TGATATAATTGAATTAAGGTTTTGCTTATGAGATTGTGCCGTCATATCCTGTAAATGAAT
AACATCATATTTCTTTTTTTCTGTCTTTCTGCAAGTAGGTCAAGCTCTGAAATTACTG
TATCACATATTGATTATATATTCTAACAATGTTGATGTTGATCCACATGGTGTATTGTT
CAAATTAGAAAAACAGTGCCTTCGTGAGTAGGATTAACCTGTCTCAGGACACTTGG
CTTGGCAAACTAGTGTGGGTAGAATTTCTGCTCCTAATCAAACCTGGCTGGGTGCCCT
TGTGCACTGAACAACCTGAACAACCTGATGCAATGGCCCTGTCTTCTGAAATCCTATAGA
TTAAGCATTGCAACAAGAAATCTAAGAATGACATACCTCATGATACTTTTTACGGTTTT
CCCTGAATTGCATGTGCAGGACCTCCAGTTGACAGTTCCACAGCCACAATTTCTCCACA
GCGCTGCACGAGGAGGCAACGTGGAAAGAACACATTGGCCAACCTCAGCTCTTCTCT
TATATTGACCGAGTAATTCCTGGGAGTGCAACTGTAACGCTTGGCATCATCATTGAGCCT
ATCCAGGTCAACTGTAAGCAAATACATGCACTGTGTAAGCAAACACAACAGTAAGCACAA
TTGCTCAGCATGTGTTTTGGGGAATAGAGTGGAAGATTCTGTGAGGACAACTTTAATCC
TGGGATTAAATCCATCCTCAGGCTCTCAATTACTGGGAACCTAAAGTAAAAGACCTAATT
TCTCTCTTGATCCCATCTACCTTTGAAAAATAACAACAAAAACAAAAACAAAAACC
TTGGCTAAAATTCATGTCTTGGTTCAAAATTCCTTGCACTCTTTTTATGCAAAAAAGTT
CAGAGGAAAAGAAAAATGAAACATTTGGCAAAATTAATAAATTGCCCTTTATCTTCAAAAAA
CAAAATAAACTGCTTCTTCATTCTAAATAAATCCTGACTCTAAGCCATCACTTAAATATA
ATTTAAATTATATCTAGGACACATTAGAGACAACCACTGTATTATATTTACCTGGCAGTG
AAATTTACACATAAATGTAATAAAGAAAAGGGTTTTAGTTTTATTTTTAACTTCAAAGT
TAATCAACATGTTAGAGAAAATGATTTTTTGTGTGTGAATATGGTGCATATTTGTGCAC
TGGGTCTAGATGGCAGTATAGCATAGTGGTTAAGTGCTTTGGAGCCGAGTAAAGTTAC
AACTGCCTTCAATCACATCATGGCTATGTTTCTTTTAGTTGGAAAGTTATTTAATGGCT
CTGCACTCAATGCCAGCTATTCTTACTAATCTTGATTCACTGAACTCTCCCCCTCTCACA
TTCTTTTCTTTTAAAGTCAAAATGGACAAAATTTAAAAATATACTACGTCTTTCTCCCTCT
TTATGTTTTCTTCTTTGGACATTATATAAATGATTATCAAGGGATATGGTCAGTGGGACT
TCTATGAACTAAAAGCACCACAATATTTTCAAAGCTAAGTCATAAATATTTACTTGAATT
CAATTCGCATTATTGTGTTACATAAATTAGATTTGATTTTCGTGTTCTCCAAGTCCATCCA
ATTATCCATCCATTTGGCAGCCAAATCCTCTTATTGGTGCATACATAGATATCTTC
TCACAAGAACACAGTTAGAAGCAAGTCTTTTGTCTGCACAGAACCCACCACCTTTCCC
AATGCACGCCCTTGTGTGCACCAGGCCACCACAGAAGGAATCCCATGAATGTTAGGCAGT
TATACCATGTTCAATTCTAACTCTAGCCCTGGTAAATGTCTTAGGCTTAACTGAACT
GCTTAAGGAGAAGGAAGAAAAAAGCCTTTGGGAGCAGAGGTCAAAGTCAAGGACAGTG
TCAGTAGAATGCATCTAAGGAAAATTGCAAGAACAGTGTGGACCATGAGTCCACTAACC
GCATCCCTGTTTTGTCTTTTTATATCTACAACCTGTGCTTTGATTCTCAAGGGCAGGGATTT

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

ATCCCCAATTGTCATACGCTGAGCCCAGGATACTGTGGGAACTAAATAAATGCCAAATAA
TCACTGTGCATAATTTATTGCTGTTTTTTTCTTCCTTTGTAAAGACATGTCATAACTTTA
GCTATAATCAATCCAGATTTAAATTGAATGAACTCTGTCTTGGGAAAGGAACCATTTTAT
TGCTGCTAAAAGCCCTGAAAAAATACTGCAAGCAAAGCAGAAAGGATTACATAAATGGATTA
AGAAAGTGCAGCAATCATCCATTACAAGTGTACCTTGAGATAAAATACTGATGCTGGCCTG
TGCATTCTGGGCTTGTGAACAGGGCTGATGCAGAGGGCCAAATCTGCTGCTTGTGATAA
TGGGAAGATATTTGTCTAACACGGAGGACTAAAGAGGTACAATTAGCAGGAAGGGATCGT
GATGTGTGAGAAGGCAGAGAGGGACAGGACGCAAGCGGCAACACTGAGTCTCTGCTTTTCG
TTCAGCCCTCTGCTGAACCTATGTCATAACCTTACAGATATGACTACATTTAGTTGTTCCC
CACCCATTTCAAGTGAGTATTATTTTGCTAAGAGTTAACAGTAAATGGTAAAACTGGAATT
GAAGCCTTTCCCCCTTTTGTCTCCATAGTCTCTGTCACTTTAAGCAGAATAGAGGGATGGT
CAGTAAGCCATTAGTTGAAGGAAGAAGACCAGTGTATGAGGGCTGTACTGGACTTTCCCT
GTCTGATGTTGAACCAGGGTTGGGTGGTGAATGGCCACATCCTTATCTTCAGAAGACACC
CAAGCCAAAGTACAGCATGCCTTCCCTATAGGAATTCCAATAAACTCCAAAGTGCCCTCC
ACAAACCAGGAGAAGGCATGTAAGCCTCATTTCTATTTGAAACCAAACCTCCTCACATTGTA
TAACTCTTTGACTAGCATTGGTTACTTTTCACTTTCTTACTCACCATGAGGAGATTACAAAT
TCTGTCACTTTTATGCATCAATTCACTAATGTCACTAAGAACTGTGACCCAGTGTGACTG
CTGTAGAATCCACCTGCCCTACATATCAGGCGATTGAAATCACCGAGGCACTCAATTTAG
AATTAGAAGTGAAGGTTGTTCGACTGGATGCAAAATACTTCAAAGCGTGATTGCAGAC
TTTCTGCAGGGGATATTGAGTTTCCAGTCAAAGAAGACTGATAAGCCAGTGCAGCAGAGC
AGGGAGCCAGACAGAGGCTGAGCAGCAATTAAGGTTTCTGGTGCTTGGGTTAGAAGGAGA
ATTCTAGAGTGAGACTTCTGGGAGGGAGACAGGCCAAGTGGTGTCAACCGAGAGAGAGGG
GCTGCCCTTTCTGTTCTGTCATCAGTCTCCATAGCTGAACATTCAAGAGCATCAAGGGTGT
TCCAAATTTAATTCAAGGGCCCCGTAATTTCAACTAACGTCCCTTATTTATCCCAGTCT
ACATCATGGAATAATCCCTCTTTTTCTGTCACTTCTGAATTTTAGTTTGTCTTTATCAAAG
CTCTTTGACATCAATTAAGTTTTTTGGGCCCTCATAACGGCAGTTTCTGGAAGGGCTGGG
AAGGCTCCCCGTCTCTCAAAAACCTTCTTTGGACATACTGGTAAAGTGTGGGGTGTGCA
CAGTGACACAGGGGCGCACTTTTCTTCTGACCCCTGCCTCTAGGTGCAGCCTAATCACA
GAGGTGCTTTGAGATGTAATAAGAATGCTGGGATCTGAGCTCTAGAATGGGGATGTGAAT
GCACACCCCTTTTATAGTTTCTAATATTAGGTGTGGGAGGAAGGAGAGTCAACCTGTTCTT
TTTGATCCTGGCCATCTAGACTACTAAGTGCTGTGTTTCATCATGAGTGCATGAAATACTC
ATAAATGCTATGAATTACATACAATCATTACTCTTACTTTTGCAGTTGAGGAAACTGAGC
CTTTCAAAGTTAAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
GTAGAGCCAGAGTTGAATCCAGGCCCTTTCATGCCTTTTGTACATTATGAGCCCAGTGT
AAAGATTTGAAGGAGGGTTCAGAGAGCAGCTTCCACAAGCGATGCTGAAGTCTTCCCACC
AGAGGACCACTTGTGTGCAATGACCTGGCTTCTAGGTGCACCCATAGGCACTAACCT
GTCATTTTCACAATGTCATTAATACACATACCCAAGCACGAGCAGCTGTGTATCCTCATC
CACCTATTCAGTTCTGACTAATAAAACCAGGCAGCTATTCAGAGTCTACTTTTGTCTTTGG
TATTGTGATAAAACATTTTCGTCCTTTTATCTGAACCCGAGTGGACTTTGGCCAATGCACAC
TCTTCTTTTGTCTCTCATTTCTCCTTGAACCCCTTTCTTGCCTCATCTAGGAGCTGAATGG
ATTGATAGCAATGATTTGTCTCAGTGGGAACAAATGCCAGGATCCTAAAGATTCTAGTT
TCTGTAGGACAGATAAAACCAATATAATATTTACATACTTGGCAAGCACACATGGGTG
GTGATCAGCTTAAGATAAAAAATTTAAAAATCTATTCCTTATTAAAGTGAAACATATTTAT
TATTTTATACAAGTTACCTGTTCTGGTGGACAGAAGTCAAAATGTGCCTTACTGGGTCAA
AATCAAGGTATGTAAAGGTAAGTCTGTTTTTGTGAATACAAAAATTCATTATGAAACAAA
ATAATCTCCTCACTTTGAACATGGGAAAGCACCTTTGTTCTCCACATGCACAATAATTTA
CTGCTTGTTTTTGGTCAACAAGAGGATGGTTTTCTTGAGTCAGTCTCTAAACATTAAAGTTA
TCAGTTGGCTTTTCACTGTGCTGGCCATTTCATCAATCTTTATTTCTGATAGTAGAATCACAT
GCTTTTTTCTCTTGATCATCCACAGACTGTCAGAGAATCCAATTAAACTGCTGCATGTCC
TTGAAGACAGTCAGTAATCTTTTTCTTTTTCTTTTTCTTTTTTTGAGACAGAGTCTT
GCTCTGCTGCCAGGCTGGAGTGCAGTGGCATGATTTCACTCACCATAACTTCCGCCTC
CCAGGTTTAAGCAATTCTCCCGCCTCAGCCTTCTGAGTAGCTGGGACTACAGGTGTGTGC
CACATGCCAGCTAATTTTTGTATTTTCATTAGAGATAGAGTTTACCATGTTGGCACCA
GTAATCATTTTATTGTTGTACAAGGTCTAGCCTTTATAATGCTAAAGTTCTTTATAAAC
TCTTAGCAGGCATGCAGCAGCAGCCTACATACTAGTGCATATTTACCACATGGTGTATGGC
ATGTGGAGAATGAGCACTCAAAGTAGGGCCATGATGTTTTTGA AAAACCATGCTCTTGTTCT
TTCTGTTTTGCTTTCTCTGGCATCTCAGCTTGCTATATTATTTTAATCTTCTAGAGCTCA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

AGTCAAAAGGCATATTTCACTCTACATCTCCTATGACTCCTTCTCTATGTACTCACCACAC
TTTTTCTGCCCTCTATTATAGCATTTACTTTCATAGACCTTGTTTTGTTTTCTTTGCAGAT
AGTCCAACCTTACAATGGTTCAACATGAGTTTTTTTTTTTTTACTTTTTTGATGGTGCTTTC
AGCTGTGTATGTTAATGGTGAGCACCATAAAACCATTTCTGTTTTTCACTTTCAGTAAAG
TATTCATAAAATTATATGAGCTATTAAACCTCTATTATAAAATAGGCTTTGTGGTAGATA
ATACTGACCAACTATAGGCTGATATATGCGTTCTGAACATGTTTAAGGTAGGCTAGGCTT
AGCTACGAAGTTCAGTAGGCTATGTGTATTGAATGTGGTTTGACTTAACTATATTTTCCA
ATTACAATGTGTTTATTGGGATGTAAACCCATCAAAAGTTAAGAAGCATCCATATTCAGG
TGCTTCTCTTCTATTGAACTAAAAAAGTTGTTGGAGGTAGAGAGCTAATCATCTTTTTAT
CCCTTCTCTTCTAGTACTTTGTCCCTAATGTAAGGACCAAATGCAGCTCATTCAGCTCA
CTTCTAAAAACCATTTGACTTCAACGATTAGGAGCAAGGTGCTGCTGAAGGAAGCACAA
TCGTCTTTGTGCCAGATTCCCTTTGCAAGAAATAGACACTCAATAATTACTACCTGAATG
AATCAGATAAAACCATGGATTTTTTTCAGTTTAATTTTTTATCACAAAGGTGAAACAATTCC
AACTATCTTTTCAGGAATGCATAGAGTATTATTAATACTTTATATGTTGAATGTTTGGATC
TTATCTCAAGACTAGTCTAGGATGTGGTGTAGTGCAGAAAGTGTAGGTTAGAATTAAGAA
ACTAGGATTTAGAACCATCTCTGCCAAGGATGAACCAAGATGCATGAGATGACTGGCACC
AAGATCTGTGAAATATAGGAGTGGCACACATGGTAGAGCATGGTCTTGAGAAGACTGGTT
CTCAGATGCAACATCAAGGTGCTAAGGAACATTGCCTTTGTCAAATGGGAGACAGTGCC
ATAATCTGTTATGAAAAGTGCTAAACATTTCCCTTTAATGTTGCAGCACCAGGTTTATGA
TGCTTTGGGGTGGGGTGGGAGGTGAGGGGACTATAGGAATAATTGTATTATTTTTGTAA
ACAACCAGGGGCTCTCAGCCTTTCTCAAGGTTTATAAGTTTACCAGGAGTCTTCATCCTT
ATGAATATTACAACCTCACTGTGATTTTTCAGAATCCCTCTTCTATAAAATCAGCACATAAT
TATGAAAACCTGTTTTTACATTTCAATCCCAATGCAGAAAGTGCAAATAAATCTGGAAAGGA
CAATTATTTTATAAGTTTTAATGAGGGACTATCAGGATACTTTAACTTAAAAATTATGAAA
GACACTGAATTTTGTGACTCCACATGTGGGTTTAACATCTAAAGAATAAAAAACATTTACC
ACTTGTTTTTTACACAAATTAACCTAAATCACAGCAAAAGATTCTTAATCAGACAATTCCT
GATTTTCAAGGTTTAGAAGCATTGTGAAGCAACGAGTGAATGCAAACTTACTGAGCAACT
ACTATGTATCCAGCACTTTGTGAGATGCTTTTCAAAAATCAAATTTTCAATTAATTCAC
AGCAAAACCTCATCTTAGATATTACAATTTCCCATTTGGATGGATGTAGAACTGAGTTGCA
GAGGACTTAAGTGACTTGCTTAAAGCCACTGGACGTAAGTAGATATTCAGCACATATGTA
CTGGAATAATCAATGGATGACTGAATGAATATCTGGAAGGCACTGAATTAGCCATGAAAGC
AGTAACACCATCATAAAGCCATTCTCTGAACCCCTGGGAATTCATGACCTGGGGTGATGTA
TAAAGCATATCTATGAAATGAAAAATTTTTTAAAAAACAAGTTAGGAGATGAAGGTCTTA
AAATTTCTTGTCTCATTTACTAGACTATGAATCCCTGAGCTTTAAACACCATTGAGTTG
CTCACAGCATGAGCTTTGCAGCCAACAAGGCCAAGTTCATGTGATGATTCTCTATTTTCT
AGCTCTCACAGAACTTTCTGAATTTTTTCTTTTACTATGCAAAATGGGAATCACAGTA
GCTCCTGCCAAGGAACCTGTAAAGATTAAATGAGTTTGTAGTGTGTTGAAGCACAGGTCTA
AATCCTACCTCTCAGTAGGTATGCCACATTATTACTGAAGTGAGTAACACAGATCAAAG
AAGGAAGTTTAAATCAGAAGCACCTAAAACTAGAAGGAGCTGAGATTCTGTAATTGGTGA
TTACTAAGGAATAGAAGGCCATGCCCACCTTGAACATAAATAAAAAAGCCAATGATCAGG
TGTCTTCAGTTTGATTACATCAGACGTAACACTATCCTTTGCAAGTGTATTAGTCCATTC
TCATACTGAGATAAAGAAAACACCCAAGACTGGGTAGTTTATAAAGAAAAAGAGGTTTAATG
GCCTCACAGTTATGCATTGCCTGGGAGGCCAAGGAACTTACAATCATGGCAGAAAGCG
AAGGAGAAGCAAAGGCACGTCTTACATGGTGGCAGGAGGGAGAGCATGTGTGCAAGTGTA
GGGGAACCTGCCCTTTATAAAATCATCAGATCTTGTGCCACTCACTCACTATCACAGAAT
AGCATGGGGAAAACCATCCCATGATTCAATTATCTCCATCTTGTCTTCTCCCTTGACATG
TGGGGATTATGGGGATTATGGGGATTGCAATTCAAGATGAGATTTGGGTGGGGACACAAT
GCATAACTATATTAGCAAGTAAAGACACTCAGTGAGTTGGATCTCTCAGCACCAGCAGGG
CCTACAAGCATATATGCTCTAGGAGCAGTGTAAGTGTCTCCTTAAGTCTAGTGGGGCATAA
AGGAAAAACAATCCCATAATAATTTTCCAATGCTTCACAGAAAAAATTAGCACAGCAAAC
ACAAGGAACACACATTTCTCTTTAACAAGAGTAATTGCAGTGGAATAATGCACATTTGTTC
ATCCGACGCTAAAAGTTACCTATGGCTTTTCCACTGTCAACTGGATTTTTCTTATTGATTT
GCATTTGAATGACATGCCATAGATGAGGGGAATAACTTTGATAATGAGGGTGGGGTTAGGA
TATCCACAAAGACGGACAACCTGCGTCTAGGATGAAAGCAGAGGTGGCACAGGCACCAGGA
GGAAAAACAACCTGAAAGTTGTCCCACTGCTGAGATTTTCTTAAAAATTTTCATGTGTGTG
CCCTCATAGACACACACAATATGATAAACAACAATATGGTTTTTATGAATGCTTGCTGGC
AAACAGAGTAAGTGAGGCAGCTAGATACTTACAAATCATATGAGTCATAAACAGTGGCAA

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500